



L'évaluation des savoirs scientifiques : modalités et enjeux

Muriel Lefebvre

► To cite this version:

Muriel Lefebvre. L'évaluation des savoirs scientifiques : modalités et enjeux. La publication scientifique : analyses et perspectives, Hermès, pp.299-316, 2008. sic_00429967

HAL Id: sic_00429967

https://archivesic.ccsd.cnrs.fr/sic_00429967

Submitted on 5 Nov 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

L'évaluation des savoirs scientifiques : modalités et enjeux

Muriel Lefebvre
Université de Toulouse – LERASS
Muriel.lefebvre@univ-tlse1.fr

In Schöpfel, J. (dir). 2008, *La publication scientifique : analyses et perspectives*, Paris, Hermès.

La rédaction de publications constitue une des activités essentielles de tout chercheur. Ces publications ont souvent plusieurs rôles : diffusion des connaissances, pour s'assurer la paternité d'un résultat, support de conservation d'une démarche ou encore instrument de validation des résultats proposés. C'est cette dernière dimension qui va nous intéresser plus spécifiquement. La recherche scientifique et les activités d'un chercheur sont en effet essentiellement évaluées depuis de nombreuses années à partir des publications d'articles [MER 73], [CAL 93], et chaque chercheur s'efforce de publier dans des revues reconnues pour valoriser ensuite son travail de recherche.

Mais comment ce système a-t-il été mis en place ? Quelles sont les procédures actuellement en vigueur pour évaluer le contenu des publications ? Après une première partie exposant la genèse et le rôle des revues dans la communication scientifique, nous exposerons les différents systèmes d'évaluation qui coexistent aujourd'hui, renvoyant selon les cas à une expertise de contenu ou à une mesure de l'impact d'un article. Ces deux systèmes font actuellement l'objet de nombreuses critiques par les chercheurs de toutes disciplines et plusieurs initiatives numériques ont récemment été développées, présentées comme des alternatives aux dispositifs traditionnels d'évaluation. C'est à leur explicitation que sera consacrée la troisième partie de notre article.

5.1. La revue au cœur de la communication scientifique validée

2 Publications scientifiques

Les premières revues scientifiques sont apparues au XVIII^e siècle avec la création des sociétés savantes¹. En France, le *Journal des Sçavans* a été créé en 1665 par l'Académie de Sciences ; en Angleterre, il s'agit de la revue *Philosophical Transactions* éditée la même année par la Royal Society. Ces revues, qui ont connu un succès immédiat, du fait de leur petit format, facile à reproduire, et assurant une diffusion rapide et régulière des informations scientifiques, ont joué un rôle important dans le développement et le fonctionnement de la science moderne, tant comme support de diffusion et de discussion de la science que comme instance de formalisation des pratiques scientifiques.

En effet, les revues, en structurant les échanges épistolaires soutenant jusque-là la diffusion et la discussion des connaissances nouvelles, ont contribué à rendre publics et collectifs les débats scientifiques : en publiant dans une revue, un auteur s'exposait désormais au regard de l'ensemble de la communauté des chercheurs et non plus seulement à ses collègues proches. Les revues ont donc constitué, depuis leur création, un espace public de discussion critique de la science, pour reprendre une expression de Boure & Surau [BOU 95], qui est progressivement devenu un espace de légitimation des savoirs nouveaux. Le développement des revues s'est accompagné d'une transformation du régime de la preuve en science : alors que jusqu'au XVIII^e siècle, seuls quelques témoins étaient présents aux expériences personnelles et souvent théâtralisées conduites par les savants (la renommée de ces derniers assurant la légitimité de leurs travaux), on assiste progressivement à la mise en place d'une technologie littéraire dépersonnalisant les comptes rendus d'expérience afin d'assurer un caractère public et diffusable à l'identique aux publications scientifiques [LIC 96]. L'écriture des publications s'est ensuite progressivement codifiée parallèlement aux tentatives de structuration et d'organisation des disciplines. Un format de présentation des recherches, le format IMRAD², a ainsi graduellement été mis en place dans les disciplines expérimentales. A partir de la fin du XIX^e siècle, on voit également apparaître des manuels de style, propres à chaque discipline, qui décrivent les « bonnes manières » et les règles formelles à adopter pour rédiger un article.

Depuis leur origine, les revues sont donc au cœur des processus de normalisation et d'évaluation de la recherche scientifique. Elles sont devenues aujourd'hui le principal support de validation d'une recherche mais également un instrument institutionnel d'évaluation des chercheurs et de leur carrière.

5.2. Les pratiques traditionnelles d'évaluation des publications

¹ Sur l'histoire des revues scientifiques, on pourra se reporter à l'ouvrage de Solla Price [SOL 63].

² Pour "Introduction, Material & Methods, Results And Discussion".

Deux systèmes d'évaluation sont actuellement principalement utilisés par une majorité de disciplines pour évaluer un article : un système d'évaluation de contenu, en amont de la publication (système des rapporteurs) et un système d'évaluation de l'impact d'un article, réalisé en aval de la publication (système bibliométrique). Il convient de noter cependant que les utilisations de ces deux systèmes diffèrent d'une discipline à l'autre, en fonction de la culture éditoriale de la communauté considérée, de son objet de recherche, de ses méthodes et de son fonctionnement organisationnel et communicationnel [KLI 00]. L'article constitue la publication de référence dans la plupart des disciplines dites Sciences, Technique et Médecine (STM), mais ce n'est pas toujours le cas en Sciences Humaines et Sociales (SHS) où la monographie joue un rôle fondamental. Toutes les publications n'ont donc pas le même statut et ne sont pas évaluées selon les mêmes critères dans toutes les disciplines [MER 73]. Par ailleurs dans certaines disciplines des sciences humaines et sociales comme la psychologie ou l'économie, on rencontre des pratiques éditoriales (et en particulier des pratiques d'évaluation) proches de celles des disciplines sciences, techniques et médecine ; tandis qu'à l'inverse, une discipline comme les mathématiques comporte de nombreux points communs avec les sciences humaines et sociales.

5.2.1. *Le système des rapporteurs ou l'évaluation du contenu par les pairs*

Avant d'examiner plus en détails le système d'évaluation par rapporteurs (*referees*), il convient au préalable de rappeler le fonctionnement d'une revue. Deux personnes sont responsables du dispositif éditorial : l'éditeur (en anglais *editor*) et le diffuseur (*publisher*). L'éditeur scientifique, généralement un chercheur, gère la création, le contenu. Il s'appuie sur un comité scientifique et sur un comité de rédaction, composés, pour le premier, de chercheurs reconnus dans le champ et légitimant le contenu éditorial de la revue et pour le second, de chercheurs s'occupant de la gestion scientifique quotidienne de la revue (contact avec les auteurs, les rapporteurs...). Le diffuseur, lui, a la responsabilité de la mise en forme définitive du prototype, sa reproduction, sa diffusion, ce qui suppose des savoir-faire que maîtrisent rarement les auteurs, plus préoccupés par le contenu que par la forme de la publication. Il s'occupe de la production (mise en page, design page de couverture, impression), du marketing (meilleure visibilité de la revue, catalogues...) et de la distribution (gestion des abonnements, diffusion en librairie...). C'est un métier à part entière, qui la plupart du temps ne concerne pas les chercheurs. L'organisation des modalités d'évaluation incombe donc dans l'intégralité à l'éditeur scientifique.

Le dispositif d'évaluation par rapporteurs (ou évaluation par les pairs, le *peer-reviewing*) a été organisé et systématisé avec l'apparition des comités de lecture, après 1945, dans un contexte d'essor considérable du nombre de publications,

4 Publications scientifiques

celles-ci étant par ailleurs de plus en plus spécialisées [BUR 90]. Les éditeurs ont progressivement mis en place plusieurs filtres : celui de l'éditeur qui sélectionne un article pour son intérêt thématique, puis celui des rapporteurs qui vérifient la qualité technique de la publication. La gestion de ce dispositif devient alors une responsabilité pour l'éditeur et la renommée d'une revue s'est petit à petit construite sur la rigueur de son système d'évaluation [BOU 95].

Le fonctionnement du système d'évaluation par les pairs comporte plusieurs étapes : l'article est envoyé par l'auteur à l'éditeur soit spontanément, soit en réponse à un appel à contribution sur un thème précis, soit en réponse à une sollicitation directe de l'éditeur³. Avant soumission, l'auteur s'est référé aux *instructions aux auteurs* publiées par la plupart des revues, qui donnent généralement une feuille de style mais également des instructions sur la structure des articles, la place et la taille des images, etc. L'article soumis est ensuite examiné par le comité de rédaction qui demande à des scientifiques (généralement deux) connus dans le champ⁴ d'évaluer la qualité du document proposé, sur la base :

- d'une évaluation scientifique : intérêt de la recherche, caractère innovant, solidité de la méthodologie, adéquation avec la ligne éditoriale de la revue, points forts, points faibles...

- d'une évaluation de la forme : lisibilité, qualité de l'expression, notes et bibliographie, longueur du texte.

Pour qu'un article soumis à publication dans une revue soit accepté, il doit comporter des résultats originaux n'ayant pas été publiés ailleurs. Chaque contribution ne peut donc être publiée qu'une seule fois. Une grille de lecture est en principe fournie aux rapporteurs permettant d'homogénéiser les critères d'expertise. Sur cette base, après lecture de l'article, les rapporteurs font les recommandations suivantes : la publication en l'état, des modifications (légères ou importantes), le refus de l'article. Si les deux rapporteurs sont d'accord, l'éditeur suit leurs recommandations, dans le cas contraire, il décide lui-même ou fait appel à un troisième rapporteur. En cas de modifications, les demandes sont transmises à l'auteur par l'intermédiaire de l'éditeur qui en fait une synthèse. Une fois l'article corrigé, celui-ci est renvoyé à l'éditeur accompagné d'une lettre expliquant comment les demandes de modification ont été prises en compte. La procédure peut être dite « en aveugle » (l'auteur ou les rapporteurs ont été anonymisés), « en double

³ En sciences expérimentales, les auteurs procèdent généralement par soumission directe alors qu'en sciences humaines et sociales, les revues fonctionnent plus volontiers à partir de numéros thématiques et donc d'appels à contribution.

⁴ Dans certaines disciplines, les auteurs suggèrent des noms de rapporteurs pour leurs articles (en géophysique et en médecine, par exemple). Dans d'autres, l'éditeur choisit un rapporteur parmi les auteurs cités dans la bibliographie.

aveugle » (l'auteur et les rapporteurs ont été anonymisés : l'auteur ne connaît pas le nom de ses rapporteurs et ces derniers ne savent pas qui est l'auteur de l'article évalué) ou encore publique (l'auteur et les rapporteurs sont explicitement mentionnés).⁵ C'est le passage d'un article par ce dispositif éditorial propre aux revues « à comité de lecture » qui assure ensuite la validation de son contenu et légitime sa publication.⁶

Le rapport de confiance que l'auteur (et les lecteurs) entretiennent avec le fonctionnement éditorial de la revue est ici fondamental. Il repose sur la relation entretenue avec l'éditeur, dont le rôle de modérateur/filtre est généralement présenté par les chercheurs comme essentiel (voire excessif puisqu'il peut également décider de manière parfois arbitraire du devenir d'un article). C'est l'éditeur qui permet que le système des rapporteurs fonctionne, et c'est lui qui en porte la responsabilité. Il choisit les rapporteurs en fonction de leur qualité scientifique, de leur intégrité et de leur capacité à émettre des commentaires constructifs. Il s'assure théoriquement de l'indépendance des rapporteurs vis-à-vis des auteurs, en déjouant les réseaux de chercheurs qui fonctionnent parfois selon le système des « renvois d'ascenseurs ». Il s'assure également du fait que l'auteur et le rapporteur ne sont pas en compétition sur le sujet traité. Mais pour une petite revue, surtout si elle est jeune, le risque existe d'avoir moins d'exigence et de rigueur, que ce soit en confiant le travail de rapporteurs uniquement au comité éditorial de la revue, en fonctionnant par sollicitations auprès des auteurs (les articles commandés sont ensuite rarement refusés) ou enfin, par « copinage ».

Depuis une vingtaine d'année, ce mode d'évaluation par les pairs est souvent critiqué par les chercheurs⁷, l'expertise d'un article étant un processus particulièrement long qui, selon les disciplines, peut prendre de six mois à deux ans,

⁵ Dans les disciplines des sciences expérimentales, le système des rapporteurs fonctionne en général avec le seul anonymat des rapporteurs (parfois levé par des rapporteurs décidant de signer ponctuellement leur évaluation pour montrer qu'ils assument leurs prises de position (positives ou négatives)). Les auteurs sont parfois anonymisés en sciences humaines et sociales mais ce n'est pas toujours le cas. L'anonymat des rapporteurs est souvent considéré comme fondamental (en terme de liberté pour le rapporteur mais également pour l'auteur qui ainsi ne se sent pas redevable, par exemple si le rapport est extrêmement élogieux) et celui des auteurs peu pertinent (on parvient presque toujours à retrouver l'école de pensée, le laboratoire, voire l'auteur de l'article, surtout lorsque celui-ci a déposé son document sur un serveur de pré-publication). Finalement, le travail d'évaluation s'effectue dans une communauté très restreinte où « tout le monde connaît tout le monde ».

⁶ Il existe cependant de fortes disparités dans l'utilisation de ce système, d'une discipline à l'autre et selon les revues considérées. Le statut de l'article et de la monographie n'est pas le même en STM et en SHS et les taux moyens d'acceptation d'un article sont très différents en STM (autour de 80%) ou en SHS (autour de 20%).

⁷ Voir notre enquête de 2006 [LEF 06].

6 Publications scientifiques

voire plus (comme en mathématiques ou en économie), et qui est peu valorisé⁸, ce qui explique peut-être en partie l'absence de sérieux de certains rapports, et leur style parfois peu courtois. C'est finalement un processus très coûteux en temps, un article refusé dans une revue finissant souvent par être publié ailleurs après une seconde évaluation. L'anonymat est par ailleurs rarement respecté, et il est facile d'identifier l'auteur (analyse de la bibliographie, auto-citations, références à une école de pensée...) voire les évaluateurs (dans certaines disciplines, il n'existe que 2 ou 3 spécialistes d'un thème dans le monde), lorsque l'article n'est pas le résultat d'un système de « copinage ». L'anonymat favoriserait de plus le désengagement des rapporteurs, qui ne sont pas forcément aussi spécialistes que l'auteur de l'article et ne peuvent vérifier l'intégralité des résultats présentés (erreur, fraude, plagiat ne sont pas exceptionnels⁹). Les rapporteurs valoriseraient par ailleurs souvent davantage les recherches classiques que les innovations et renforceraient ainsi le courant intellectuel dominant. Enfin, en l'absence fréquente de grille de lecture et donc de critères précis sur lesquels s'appuyer, l'évaluation comporte toujours une importante dimension subjective (un jeune chercheur n'évalue pas forcément de la même manière qu'un chercheur senior), souvent dénoncée, surtout lorsqu'il n'existe pas de droit de réponse systématique de l'auteur aux rapporteurs.

Malgré ces critiques et ces inconvénients, ce système continue à être utilisé et à servir de modèle de référence pour légitimer une publication, notamment parce comme l'ont bien montré Piolat et Vauclair [PIO 04], ce dispositif participe directement à la production des connaissances que ce soit par sa dimension heuristique¹⁰ ou encore par le rôle social et communautaire qu'il joue au sein des communautés de chercheurs.

5.2.2. L'évaluation par les bases de données de citations : mesurer l'impact d'une publication

Outre cette évaluation de contenu, un article peut également faire l'objet d'une évaluation de son impact supposé, après sa publication. Parmi les outils bibliographiques utilisés, on peut citer le *Science Citation Index*, créé en 1963 par Eugène Garfield, qui répertorie par discipline les citations qui sont faites d'un article

⁸ S. Harnad [HAR 00] parle à ce propos de « invisible hand », ce travail d'expertise n'étant généralement pas reconnu publiquement.

⁹ On peut citer l'affaire Sokal ou celle concernant un biologiste coréen, le professeur Hwang qui avait annoncé en 2005 avoir cloné un chien. Ce dernier exemple a donné lieu à de nombreux éditoriaux, courriers ou articles parus dans la revue *Nature* et questionnant les modalités actuelles d'évaluation des publications [GIL 06], [WAD 06].

¹⁰ Le rapporteur sélectionne les articles mais aide également les auteurs à « accoucher » de leur publication. Il arrive d'ailleurs que des collaborations auteurs/rapporteurs se mettent ensuite en place.

donné dans la littérature spécialisée (environ 3700 revues considérées comme centrales sont analysées, dans le seul domaine Sciences, Techniques et Médecine). Des outils similaires ont été développés en sciences humaines et sociales dans les années 70 : le SSCI (*Social Sciences Citation Index* analysant 1700 revues) et le AHCI (*Arts and Humanities Citation Index* analysant 1130 revues)¹¹. Initialement ces index permettaient d'accroître les connaissances statistiques sur les publications scientifiques. Ils sont aujourd'hui utilisés par les chercheurs pour trier dans le flot d'articles produits chaque jour ceux qui ont été les plus cités en partant du présupposé que s'ils sont cités, c'est qu'ils sont pertinents et intéressants.

L'utilisation de ces outils a conduit à un déplacement progressif des critères de jugement de la qualité des travaux scientifiques : du système collectif d'évaluation de contenu par les pairs, on est passé à une évaluation en aval par taux de citations que reçoit un article après sa publication. Leur succès a renforcé le rôle clé des publications et ces outils ont rapidement été utilisés pour mesurer quantitativement la productivité de la recherche scientifique et des chercheurs. Ainsi, de nombreuses institutions, dans le monde entier, y font aujourd'hui référence comme critère d'évaluation. C'est le cas en France de la plupart des sections CNU (Conseil National des Universités) lors de l'attribution des postes d'enseignants-chercheurs (les candidatures sont examinées au regard du nombre de publications produites par les candidats et de leur impact) ou encore des budgets et subventions accordés aux équipes de recherche (calculés en fonction du nombre de chercheurs dits « publiant » présents dans un laboratoire¹²). Par leur simplicité et leur rapidité d'utilisation, les outils statistiques d'évaluation ont donc une influence décisive, dans de nombreuses disciplines (en biologie, en chimie notamment) sur la carrière des chercheurs et le devenir des équipes de recherche.

De nombreux biais sont néanmoins fréquemment évoqués. Cet outil correspond aux modalités de publication des disciplines dites « sciences, techniques et médecine » (en prenant essentiellement en compte des articles de revues anglo-saxonnes) mais n'est pas approprié pour étudier les sciences humaines et sociales grandes productrices de monographies. Une grande partie de la production scientifique mondiale (rapports, ouvrages, mémoires, etc.) est ainsi laissée de côté. De même les pratiques de citations sont extrêmement variables d'une discipline à l'autre, d'un chercheur à l'autre. Des articles sont parfois cités sans avoir été lus

¹¹ <http://scientific.thomson.com/> [site visité le 10/01/2008]

¹² Sont considérés comme « publiant » les chercheurs ayant produit sur une période de 4 ans au moins :

- soit 2 ouvrages
- soit 4 articles parus dans des revues à comité de lecture
- soit 8 conférences dans des colloques internationaux

ou toute autre combinaison équivalente.

8 Publications scientifiques

(phénomène d’auteur ou de thème à la mode), tandis que certaines références peuvent être omises de manière délibérée. Par ailleurs, l’auto-citation et le « renvoi d’ascenseur » permettent à un chercheur d’avoir un bon indice de citation de manière artificielle. De plus, un auteur sait implicitement qu’il a plus de chances d’être publié s’il cite des articles de la revue où il a choisi de soumettre son document. Enfin, un article fréquemment cité n’est pas forcément un bon article : il peut s’agir d’un mauvais article très souvent critiqué.

Le prolongement de cet outil bibliométrique est le facteur d’impact qui mesure la notoriété d’une revue. Cette notoriété est calculée en fonction du nombre de citations dont les articles publiés dans une revue font en moyenne l’objet les deux années précédant celle examinée. Une revue dotée d’un fort facteur d’impact attire d’autant plus les auteurs, qui espèrent accroître la visibilité de leur recherche et ainsi mieux valoriser leurs articles.

5.3. Des alternatives numériques en développement

Dans un contexte général de crise des revues scientifiques [CHA 00], [GUE 00], [PIG 05], plusieurs nouveaux modèles de publications dits « en accès libre » ont été développés pour contourner les systèmes commerciaux¹³. Parallèlement, des modèles alternatifs d’évaluation s’appuyant sur l’auto-archivage et les revues électroniques ont été mis en place par des chercheurs militants pour un accès libre aux résultats de la recherche. Ces « nouveaux » modèles d’évaluation proposent des procédures d’évaluation à la fois pré et post-publications, plus ouvertes que celles traditionnelles et souvent publiques.

5.3.1. De nouveaux supports de diffusion et de publication

Pour présenter ces systèmes alternatifs, il est nécessaire de distinguer les deux principaux supports numériques de diffusion des publications scientifiques, à savoir les revues électroniques et les serveurs d’auto-archivage.

5.3.1.1. Les revues électroniques

Les revues électroniques se sont multipliées depuis une quinzaine d’années. Parmi les premières revues, on peut citer *Surfaces*, dirigée par J.-C. Guéron ou encore *Psychology*, revue de psychologie et de sciences cognitives dirigée par S.

¹³ L’objet de cet article n’est pas de brosser un tableau général du contexte du développement de l’Open Archive Initiative. Pour davantage de précision, le lecteur pourra se référer à d’autres contributions du même volume.

Harnad. L'avantage considérable de ce support est incontestablement le faible coût de l'édition électronique, par rapport à l'édition papier.

Pour la plupart des chercheurs, quelle que soit leur discipline, si les revues électroniques constituent l'avenir des publications scientifiques en terme d'accès à l'information, elles doivent avant tout « vieillir » pour acquérir une certaine crédibilité et une notoriété, surtout lorsqu'elles n'existent que sous forme électronique et ne bénéficient pas de la légitimité d'une version papier déjà reconnue. Les revues électroniques (en libre accès ou non) sont en effet considérées avec méfiance par une partie des chercheurs ne les utilisant pas régulièrement (et préférant le support papier) [MAH 00]. Les avantages indéniables : délais de publication plus courts (l'article est mis en ligne dès qu'il est accepté et non un an après), accès plus rapide par téléchargement, possibilité liée au support numérique d'insérer d'autres supports comme des films en astrophysique, accès plus large notamment dans les pays en voie de développement, etc., sont mis en avant. Mais une méfiance perdure : Quelle sera la pérennité de ces revues ? Y aura-t-on toujours accès dans 50 ans ? Surtout, les articles publiés dans ces revues sont-ils bien évalués selon le système traditionnel des rapporteurs ? Pourtant, les revues électroniques, qu'elles soient ou non en libre accès, adoptent généralement les mêmes procédures éditoriales que les revues dites « papier ». Elles conservent donc le modèle traditionnel d'évaluation par rapporteurs, à côté duquel certaines revues électroniques proposent un système de forum et de commentaires, tandis que d'autres ont développé des outils statistiques de mesure de l'impact d'un article (nombre de fois où un article a été téléchargé, nombre de fois où il a été cité, etc.). Nous exposerons plus en détail ces dispositifs dans la suite de cet article.

5.3.1.2. *Auto-publication et auto-archivage*

Pour diffuser rapidement des résultats nouveaux, les chercheurs pratiquent également ce que l'on appelle l'auto-publication : ils rendent accessibles sur leurs pages Web personnelles, sur le site de leur institution ou encore dans des archives ouvertes, une partie des documents formalisés qu'ils ont rédigés. Il peut s'agir de documents déjà publiés – les post-publications, ces documents qui ont une « référence de publication »¹⁴ et qui peuvent être aussi variés que des articles parus dans une revue à comité de lecture, un chapitre d'ouvrage, une communication avec actes, un rapport, une thèse, etc. – ou de pré-publications. Les pré-publications sont des documents de recherche rédigés par les chercheurs et destinés aux chercheurs, qui comportent toutes les caractéristiques formelles d'un article destiné à la publication mais qui n'ont pas été évalués selon la procédure classique d'évaluation dite « par les pairs ». Il s'agit généralement soit de la première version d'un

¹⁴ Pour reprendre la formulation proposée sur le site HAL du CCSD (Centre pour la Communication Scientifique Directe) du CNRS lorsque l'on fait une recherche par type de publication. <http://www.ccsd.cnrs.fr/> [site visité le 10/01/2008].

document que son auteur souhaite voir publier, soit d'une communication à un colloque, par exemple, dont les actes n'ont finalement pas été édités. Ces pré-publications sont des documents « publiés » (dans le sens de « rendus publics ») directement par leurs auteurs, sans recourir à la médiation d'un éditeur et sans le support d'une revue (qu'elle soit ou non papier). Pour déposer dans les réservoirs prévus à cet effet, les chercheurs doivent simplement remplir un formulaire d'identification personnelle et institutionnelle avant de renseigner les champs de métadonnées concernant le document déposé.

Le dépôt dans des archives ouvertes est le support d'auto-publication qui rencontre aujourd'hui le plus de succès et qui tend à s'institutionnaliser. Certaines archives ouvertes comportent principalement des articles parus dans des revues à comité de lecture¹⁵, d'autres plutôt des pré-publications¹⁶, selon la tradition éditoriale à laquelle est rattachée la discipline considérée. Quoi qu'il en soit, de nombreuses institutions, comme le CNRS ou les universités, incitent désormais fortement leurs chercheurs à déposer leurs pré- et leurs post-publications dans les serveurs créés à cet effet¹⁷, pour assurer leur visibilité sur la scène internationale. Cela amène pré- et post-publications à se côtoyer simultanément sur Internet. Les lecteurs peuvent ensuite télécharger librement les documents qui les intéressent.

5.3.2. Les modalités d'évaluation du libre accès : vers un nouvel espace scientifique public ?

Différents dispositifs d'évaluation sont aujourd'hui proposés, conjointement à la publication dans les revues électroniques ou dans les archives ouvertes.

5.3.2.1. L'évaluation des documents déposés dans les archives ouvertes

Même si la plupart des réservoirs d'archives comportent les deux types de documents, le système d'auto-archivage concerne essentiellement la « science en train de se faire » des pré-publications, et non la « science évaluée » des publications. Le niveau scientifique attendu est généralement celui des revues à comité de lecture¹⁸ mais aucune évaluation scientifique des documents déposés n'est à proprement parler réalisée par les pairs. Les initiateurs des serveurs d'auto-archivage sont partis du principe que tous les documents déposés étant « rendus publics », leurs auteurs devaient fonctionner par autocensure pour éviter de se

¹⁵ C'est le cas de 32% des documents déposés dans ArchiveSIC, le serveur d'auto-archivage en sciences de l'information et de la communication ou encore dans HAL, le serveur d'auto-archivage du CNRS (chiffres datant du 7 septembre 2007).

¹⁶ Comme le serveur *ArXiv*. <http://arxiv.org/>

¹⁷ Voir la plate-forme HAL du CNRS : <http://hal.archives-ouvertes.fr>

¹⁸ Voir les indications de dépôt sur le serveur du CNRS HAL

décrédibiliser publiquement. Certaines archives procèdent néanmoins à un filtrage qui, s'il ne peut s'apparenter au travail d'un éditeur, doit néanmoins être pris en compte. C'est le cas par exemple d'*ArXiv* (serveur de pré- et post-publications en mathématiques et en physique), pour lequel un système de présélection large a été mis en place. Jusqu'en 2004, ce serveur procédait à une vérification de l'adresse électronique de l'auteur pour s'assurer de son rattachement institutionnel. Depuis 2004, une évaluation rapide et automatique du contenu est également réalisée par un logiciel ainsi qu'une vérification du statut de l'auteur (celui-ci doit être parrainé par un auteur connu d'*ArXiv*). L'authentification et la reconnaissance de l'auteur remplace ici une véritable évaluation de contenu. Mais le lecteur reste finalement le seul évaluateur des documents consultés, ce qui ne l'empêche pas d'engager avec l'auteur un dialogue privé¹⁹ pour demander des précisions ou pointer d'éventuelles erreurs [PIG 03].

Ginsparg, à l'origine de cet immense réservoir en 1991, propose d'aller au-delà de ce système de filtrage, peu efficace, selon lui. Pour le futur, il propose un nouveau système d'évaluation à deux niveaux [GIN 04] :

- un premier niveau d'évaluation qui correspond au système actuel d'*ArXiv* c'est-à-dire une évaluation « institutionnelle » ou de la carrière du chercheur ;
- un deuxième niveau correspondant à une évaluation en profondeur de certains textes sélectionnés sur la base d'une mesure d'impact (statistiques de lecture, taux de citation, classement par auteur). Cette évaluation relèverait d'une évaluation traditionnelle avec rapporteurs et/ou commentaires publics.

Mais pour l'heure, la seule contrainte pour déposer dans une archive ouverte est de remplir le formulaire qui y est associé et la plupart des archives ne procèdent à aucune évaluation, même automatique, de contenu.

5.3.2.2. *Les commentaires en lignes*

Le dispositif des commentaires en ligne est récent et mis en place par quelques (rares) revues électroniques. Ce dispositif vient en général en complément du système traditionnel d'évaluation par rapporteurs, amenant une dimension interactive et publique à l'évaluation. S. Harnad en a été le précurseur puisque dès 1991, il a mis en place pour la revue qu'il dirige, *Psychology*²⁰, une évaluation par les pairs grâce à leurs commentaires publics. On parle ici de modèle ouvert de publication puisqu'il reprend le modèle traditionnel d'évaluation et l'enrichit par suppression de l'anonymat et commentaires des lecteurs.

¹⁹ Ou public, comme le propose par exemple l'archive ouverte médicale *Nature Precedings* qui publie les commentaires et les notes attribuées par les lecteurs aux pré-publications consultées. <http://precedings.nature.com/> [site visité le 10/01/2008].

²⁰ *Psychology* (<http://psycprints.ecs.soton.ac.uk>, de 1991 à 2002).

Selon les revues, les commentaires en ligne peuvent ensuite être publiés avant, après ou encore, à la place de l'évaluation par des rapporteurs. Pour l'heure, ce dernier cas de figure relève de l'exception, la légitimité des contenus publiés dépendant toujours du passage par le dispositif éditorial d'évaluation par rapporteurs. Les revues privilégient plutôt la parution de commentaires une fois que l'article a été publié. C'est le cas par exemple des forums de discussion mis en place par le portail de revues SHS *Persée*, en France²¹. Quelques revues enfin proposent des commentaires parallèlement au travail d'évaluation par rapporteurs. On peut citer le *Journal of Interactive Media in Education*²², qui fonctionne ainsi depuis 1996 ou encore la revue *Atmospheric Chemistry and Physics*²³, créée, elle, en 2001. Après une première évaluation rapide par les rapporteurs choisis par l'éditeur, l'article est rendu public dans l'espace « discussion » du site de la revue. Peuvent intervenir simultanément des chercheurs et les rapporteurs (dont les commentaires sont ou non anonymes, mais l'authentification du rapporteur rend plus crédible le contenu de son commentaire). Après quelques semaines, la discussion est close, les rapporteurs achèvent leur travail d'évaluation avant rejet définitif ou publication de l'article dans la partie « journal » du site de la revue.

Avec la mise en place de ce dispositif, on retrouve une fonction première de la revue : le débat public. Ce système permet en effet de développer l'interactivité (notamment si les rapports des évaluateurs sont publiés) et de multiplier les lecteurs et les lectures critiques en associant directement les commentaires au texte lui-même, sans décalage ni d'espace ni de temps. Les forums proposés constituent un lieu de débat de l'actualité, et organisent de véritables espaces publics de discussion et de critique de la science [BOU 95]. En ce sens, ils prolongent les revues en proposant de manière systématique un droit de réponse, souvent absent des revues de référence.

Néanmoins, les chercheurs se montrent souvent très méfiants [LEF 06] vis-à-vis de ce système. L'évaluation par des rapporteurs reste considérée comme beaucoup plus approfondie que ne le seraient des commentaires faits après parution. Surtout, les rapports rédigés par les évaluateurs participent directement à la production des savoirs en modifiant le document destiné à être publié [PIO 04], ce qui est plus rarement le cas avec le système des commentaires (surtout si celui-ci fonctionne a posteriori, une fois l'article publié), ou cela permettrait seulement d'identifier les erreurs oubliées par les rapporteurs. On peut supposer par ailleurs que d'autres enjeux (sociaux ou politiques) que ceux ayant trait au contenu de l'article peuvent se jouer dans ces formes de forums de discussion (risques de débordement, de

²¹ Chaque revue hébergée par *Persée* a été dotée d'un forum de discussion. Mais ceux-ci sont pour l'heure totalement inactifs.

²² <http://www.jime.open.ac.uk> [site visité le 10/01/2008].

²³ <http://atmospheric-chemistry-and-physics.net/index.html> [site visité le 10/01/2008].

règlements de compte, de défoulement collectif, etc.) [ROS 00]. La question du contenu scientifique des commentaires et de leur qualité doit en effet être abordée : qui va prendre le temps de lire des articles non évalués, de rédiger des commentaires, c'est-à-dire finalement de s'investir dans ce nouvel espace de dialogue, dans un univers où ce qui prime pour la carrière d'un chercheur est le nombre d'articles qu'il aura publiés dans des revues à comité de lecture ? Tous les articles doivent-ils par ailleurs faire l'objet de commentaires²⁴ et quelles seront à terme les pratiques de lectures de ces commentaires ? Autant de questions qui restent pour le moment sans réponse, ce système étant toujours essentiellement au stade expérimental.

5.3.2.3. *Les listes de suggestions de lecture*

Dans le même ordre d'idée, la plate forme de revues en libre accès en biologie et médecine BiomedCentral²⁵ a lancé en 2002 le dispositif « Faculty of 1000 Biology », auquel est associé depuis 2006 le « Faculty of 1000 Medicine ». Il se propose d'identifier les articles les plus intéressants (en libre accès ou parus dans les revues traditionnelles) sur la base de recommandations et de commentaires publics de 1000 chercheurs internationaux reconnus. De façon à couvrir de manière exhaustive l'ensemble des domaines de la biologie (et respectivement de la médecine), cette science est divisée en facultés, elles-mêmes subdivisées en sections, dans lesquelles trois ou quatre chercheurs sont chargés de lire et commenter chaque mois les 2 ou 3 articles qu'ils ont repérés, en fonction de leur valeur scientifique et surtout indépendamment de la renommée de la revue dans laquelle ces articles sont publiés. Les commentaires soulignant les grandes lignes et l'intérêt de l'article sont ensuite publiés ainsi que la note attribuée par chaque expert (3 « *recommended* » pour un public spécialisé ; 6 « *must read* » pour des articles d'intérêt plus généraux ; 9 « *exceptional* » lorsque l'article est considéré comme faisant partie des 1% meilleurs articles publiés). L'article est ensuite qualifié selon les six caractéristiques suivantes : nouvelle découverte, avancée technique, hypothèse intéressante, confirmation importante, controverse, réfutation. Le facteur F1000 est enfin calculé sur la base d'une moyenne des notes attribuées, de la manière suivante : moyenne des deux notes les plus élevées à laquelle on ajoute la somme de toutes les notes divisée par 30, ce qui permet d'attribuer le label final « *recommended* », « *must read* » ou « *exceptional* » à l'article²⁶. Une liste de « perles » (*Hiddens Jewels*) est également publiée régulièrement de manière à faire

²⁴ Ce qui amène à poser d'autres questions : quels documents peut-on/doit-on diffuser en ligne ? Toutes les thèses méritent-elles d'être publiées en ligne, par exemple ?

²⁵ <http://www.biomedcentral.com/>

²⁶ L'exemple donné par le site de BiomedCentral est celui d'un article examiné par trois experts. Deux d'entre eux auraient qualifié l'article de « *must read* » tandis que le troisième l'aurait catégorisé « *exceptional* ». Le facteur F1000 serait alors égal à : $(6+9)/2 + (6+6+9)/30 = 8.2$. <http://www.f1000biology.com/about/system/#f1000factor> [site visité le 10/01/2008].

sortir de l'ombre des articles parus dans des revues peu connues des outils quantitatifs traditionnels d'évaluation (et notamment de l'ISI) et à faible audience.

De la même manière, certains éditeurs publient une liste des articles les plus importants (à leurs yeux ou à ceux des rapporteurs), sorte de « hit parade » des publications récentes, qui peuvent être téléchargés gratuitement (plusieurs revues de la Royal Society of Chemistry proposent ce service).

Enfin, toujours dans une logique collaborative, le site *DissectMedicine*²⁷, dont une version test a été mise en place récemment par le groupe *Nature Publishing*, propose aux lecteurs identifiés de signaler et d'attribuer une note à des articles médicaux qui leur paraissent pertinents.

Ces systèmes de suggestions de lecture comportent néanmoins quelques inconvénients, surtout si la liste proposée est rédigée par d'éminents chercheurs reconnus dans leur discipline et qui peuvent être tentés d'abuser de leur autorité : mandarinat, effets de mode, citation de collègues, d'étudiants de collègues, etc. Ils permettent cependant de sélectionner rapidement des articles peu accessibles sur la base de commentaires argumentés d'autres chercheurs, reconnus pour leur expertise.

5.3.2.4. L'évaluation « par les clics »

L'évaluation dite « par les clics » est un système qui existe aussi bien pour les revues électroniques que pour les archives ouvertes. Il propose des statistiques de consultation des résumés et des articles (combien de fois tel article a-t-il été consulté ou téléchargé) ainsi que des statistiques de citations (de combien de citations ce même article a-t-il fait l'objet). Toutes les revues et les archives ouvertes n'utilisent cependant pas ce dispositif, pour des raisons souvent déontologiques (ainsi, le responsable d'*ArXiv* refuse-t-il de les mettre en place).

Nous pouvons citer l'exemple de Citebase²⁸, développé par le physicien britannique Tim Brody, de l'université de Southampton, qui calcule le nombre de téléchargements puis le nombre de citations dont peut faire l'objet un article de physique, de mathématiques, d'informatique ou de bio-médecine déposé sur le site miroir britannique d'*ArXiv* ou sur E-Print. Sont ainsi calculés :

- l'impact d'un auteur, à partir du nombre de citations identifiées par Citebase vers les articles de l'auteur, divisé par le nombre total d'articles rédigés par ce même auteur ;

- le nombre de téléchargements dont font l'objet les articles d'un auteur divisé par le nombre total d'articles rédigés par ce même auteur ;

²⁷ <http://www.dissectmedicine.com/>

²⁸ <http://www.citebase.org/>

-ces résultats statistiques sont ensuite matérialisés par une courbe.

En France, l'archive ouverte en sciences de l'information et de la communication ArchiveSIC²⁹ met à disposition des auteurs (dans un espace leur étant réservé) les statistiques de consultation et de téléchargement dont font l'objet leurs articles (version réduite et texte intégral), ainsi que les courbes temporelles correspondantes. Les statistiques de citations ne sont par contre pas disponibles.

Ces systèmes de calcul de téléchargement et de citations, même s'ils se présentent comme des dispositifs alternatifs aux systèmes bibliométriques, notamment le *Science Citation Index* (les articles analysés sont essentiellement des pré-publications et non des publications) restent critiquables. Ils sont en effet globalement peu utilisés pour évaluer la qualité d'une publication, et servent surtout comme critère d'évaluation de la carrière et de la notoriété d'un chercheur. En mathématiques ou en sciences humaines et sociales, cet outil n'a que peu de pertinence, les articles majeurs n'étant généralement cités que plusieurs années après leur parution.

5.4. Conclusion : redéfinir les modalités d'évaluation de la carrière des chercheurs ?

Il est encore trop tôt pour évaluer sérieusement les répercussions de l'environnement numérique sur les pratiques sociales d'évaluation des savoirs scientifiques, les nouvelles pratiques étant pour l'heure essentiellement expérimentales dans la plupart des disciplines. Les processus de légitimation des connaissances scientifiques sont par ailleurs extrêmement variables d'une discipline à l'autre. Ainsi, si l'évaluation qualitative de contenu constitue un dispositif essentiel en sciences humaines et sociales et en mathématiques, il est sans doute moindre en physique des hautes énergies par exemple, où la nécessité d'un partage des données, d'une circulation des informations prime sur la peur du plagiat, comme l'ont souligné Kling & McKim [KLI 00]. Cependant, si l'organisation et le fonctionnement des communautés de chercheurs, les traditions éditoriales, le rapport au support électronique et à celui papier, rendent les usages des publications et leurs évaluations très variables d'une discipline à l'autre, on peut néanmoins conclure que la question de l'évaluation de la recherche constitue un des enjeux majeurs du développement des outils du libre accès, le problème rencontré par les chercheurs aujourd'hui n'étant pas tant celui de l'accès à l'information ou de sa publication que celui de sa valorisation et de sa visibilité.

²⁹ <http://archivesic.ccsd.cnrs.fr/>

Aujourd'hui, les revues ne constituent plus, comme au XIX^e siècle, le support privilégié à la diffusion de l'information scientifique et à l'exercice de la critique scientifique. Elles sont relayées par d'autres formes de communication comme les interventions dans les séminaires ou les colloques, les échanges électroniques, etc. Les dispositifs alternatifs que représentent les forums de discussion ou les commentaires électroniques permettent de renouveler l'espace public de mise en critique de la science, en rendant possible une interactivité plus importante entre les auteurs, les rapporteurs et les lecteurs et l'évolution continue d'un texte en intégrant progressivement les critiques qui auront été réalisées. Mais ces dispositifs sont généralement perçus comme de simples compléments interactifs au système traditionnel d'évaluation par les pairs. Le système d'évaluation par les rapporteurs, malgré ses différentes limites, reste en effet celui de référence³⁰ et les revues demeurent l'instance incontournable de légitimation des savoirs nouveaux et de valorisation des auteurs (en termes de notoriété mais également de carrière) dans la plupart des disciplines.

Rien ne permet néanmoins de prévoir quelles seront les futures pratiques d'évaluation, dans un contexte où la notion de « publication » elle-même est revisitée. En effet, sont aujourd'hui publiés aussi bien des articles validés par les pairs que des documents déposés dans les archives ouvertes, rendus publics par leurs auteurs sans évaluation préalable. De nombreux chercheurs passent désormais outre le dispositif éditorial des revues – que ce soit par militantisme économique ou par facilité d'accès –, aussi bien pour publier leurs propres documents que pour avoir accès aux publications des autres scientifiques. Cet accès direct aux articles, même s'il s'agit d'articles publiés, minore finalement le rôle des revues dans leur rôle de construction d'une communication légitime et d'un espace d'évaluation de la recherche. Ce phénomène conduit à se demander si le bouleversement éditorial induit par le développement de l'auto-publication ne va pas amener à repenser la question des modalités d'évaluation de la recherche scientifique. Peut-on imaginer qu'à terme les nouveaux systèmes d'évaluation décrits plus haut viendront de fait remplacer l'ancien système des rapporteurs, celui-ci ayant montré ses limites matérielles, les rapporteurs disponibles pour expertiser un article étant de plus en plus difficile à solliciter ?

Plusieurs questions restent donc en suspend et font actuellement l'objet de discussions enflammées. Qu'est ce qui fait la qualité d'une évaluation ? Sur quels critères l'évaluation est-elle elle-même évaluée ? Quels rôles politique mais aussi social et cognitif joue exactement l'évaluation de contenu par les rapporteurs, pour rendre ce dispositif pour l'heure incontournable dans une grande majorité de discipline ? Il nous semble nécessaire de rouvrir aujourd'hui le débat déjà ancien de

³⁰ Pour les chercheurs mais également pour les professionnels de l'information, comme le montre l'étude récente réalisée par Publishing Research Consortium [BEC 07].

l'évaluation des chercheurs et de la recherche afin d'identifier de nouvelles bases d'évaluation, développées sur d'autres critères que les revues, le nombre de publication d'un chercheur et le facteur d'impact des revues où ces publications ont été publiées. Seule une modification des procédures d'évaluation de la recherche pourrait stopper l'inflation galopante du nombre de publications actuellement diffusées. On assisterait alors à une remise en question de l'un des fondements des communautés scientifiques contemporaines.

5.5. Bibliographie

- [BEC 07] BECKETT C., INGER S., *Self-Archiving and Journal Subscriptions: Co-existence or Competition? An international survey of librarians' preferences*, Publishing Research Consortium, 2007.
- [BOU 95] BOURE R., SURAUD M.G., « Revues scientifiques, lectorat et notoriété », *Recherche en Communication*, 4, pp.37-59, 1995.
- [BUR 90] BURNHAM J.-C., « The evolution of editorial peer review », *The Journal of the American Medical Association*, vol. 263, n° 10, pp. 1323-1329, 9 mars 1990.
- [CAL 93] CALLON M., COURTIAL J.-P., PENAN H., *La scientométrie*, PUF, Coll. QSI, 1993.
- [CHA 00] CHARTRON G., SALAÜN J.-M., « Le reconstruction de l'économie politique des publications scientifiques », *Bulletin des Bibliothèques de France*, t.45, n°2, pp.32-42, 2000.
- [GIL 06] GILES J., « Journals submit to scrutiny of their peer-review process », *Nature*, n°439, p. 252, 18 janvier 2006.
- [GIN 04] GINSPARG P., « Can Peer Review be better Focused? » *Science and Technology Libraries* 22(3/4), 2004.
- [GUE 00] GUEDON J.-C., « Numériser les revues savantes : d'un commerce à l'autre », *La Recherche*, n°335, pp.78-85, octobre 2000.
- [HAR 00] HARNAD S., « The Invisible Hand of Peer Review », *Exploit Interactive*, 5, avril 2000.
- [KLI 00] KLING R., MCKIM G., « Not Just a Matter of Time: Field Differences and the Shapping of Electronic Media in Supporting Scientific Communication », *Journal of the American Society for Information Science*, 2000.
- [LEF 06] LEFEBVRE M., « Vers une remise en question des modalités traditionnelles d'évaluation des savoirs scientifiques », séminaire *Communication scientifique et valorisation de la recherche à l'heure d'Internet*, Urlist-SCD Toulouse 1, 13 avril 2006.
http://archivesic.ccsd.cnrs.fr/sic_00112197
- [LIC 96] LICOPPE C., *La formation de la pratique scientifique*, La Découverte, 1996.
- [MER 73] MERTON R.K., *The Sociology of Science. Theoretical and Empirical Investigations*, The University of Chicago Press, 1973.

18 Publications scientifiques

[MAH 00] MAHÉ A., ANDRYS C., CHARTRON G., « How French Research Scientists are Making Use of Electronic Journals », *Journal of Information Science*, 26 (5), pp. 291-302, 2000.

[PIG 03] PIGNARD N., « La publication scientifique sur Internet » in LE BŒUF C., PELISSIER N. (Dir.), *Communiquer l'information scientifique : éthique du journalisme et stratégie des organisations*, L'Harmattan, pp. 367-387, 2003.

[PIG 05] PIGNARD-CHEYNEL N., « L'édition de revues scientifiques. Une forme de marchandisation de la diffusion des connaissances », *Sciences de la Société*, n°66, pp. 181-198, octobre 2005.

[PIO 04] PIOLAT A., VAUCLAIR J., « Le processus d'expertise éditoriale avant et avec Internet. » *Pratiques Psychologiques*, 10, 255-272, 2004.

[ROS 00] ROSENTAL C., « La production de connaissances certifiées en logique : un objet d'investigation sociologique », *Cahiers internationaux de sociologie*, vol.CIX, pp. 343-374, 2000.

[SOL 63] SOLLA PRICE D.J., *Little Science, Big Science*, Columbia University Press, 1963.

[WAD 06] WADMAN M., « Journal reveals plans to fight fraud », *Nature*, n°444, pp. 658-659, 6 décembre 2006.